

Zur Kurzfristprognose mit Faktormodellen und Prognoseanpassungen

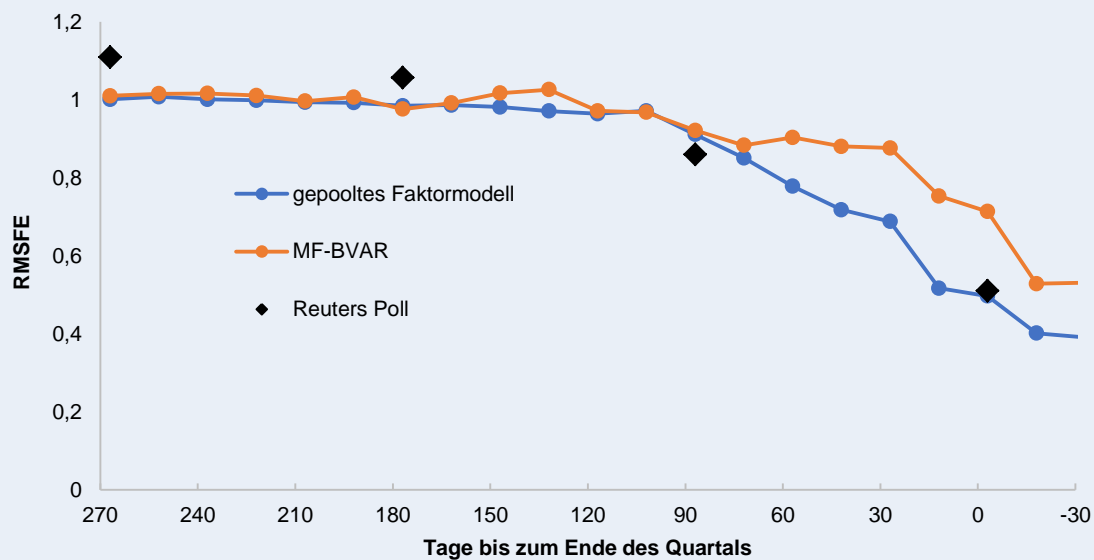
Philipp Hauber

Das Bruttoinlandsprodukt wird in Quartalsfrequenz veröffentlicht und dies mit einigem zeitlichen Abstand zum jeweiligen Quartalsende (üblicherweise 45 Tage). Zugleich ist eine Vielzahl an anderen Indikatoren in höherer Frequenz und mit geringeren Publikationsverzögerungen (z.B. Umfrage-Indikatoren oder Produktionsdaten) verfügbar, die Informationen über das Bruttoinlandsprodukt im laufenden Quartal enthalten dürften. Unter dem Begriff „Nowcasting“ werden in der Literatur Ansätze diskutiert, die mit den Besonderheiten dieser Situation für die Kurzfristprognose umgehen können.

Das IfW-Prognosezentrum verwendet für die Kurzfristprognose eine Vielzahl von Modellen. Hierzu zählen auch Faktormodelle, die in der Nowcasting-Literatur eine prominente Rolle spielen.^a Die zugrunde liegende Idee dieser Modellklasse besteht darin, den Gleichlauf von Konjunkturindikatoren anhand weniger, unbeobachteter Faktoren zu beschreiben.^b Ein Vorteil dieser Modelle ist, dass sich durch die so erzielte Dimensionsreduktion innerhalb eines Modells eine Vielzahl von Variablen berücksichtigen lassen. Darüber hinaus erlaubt die Faktorstruktur eine Darstellung als Zustandsraummodell, wodurch die unterschiedlichen Frequenzen der Indikatoren sowie fehlende Beobachtungen am aktuellen Rand handhabbar werden. Dadurch können tagesaktuelle Informationen innerhalb eines einheitlichen Modellrahmens in die Prognose miteinbezogen und ihr Einfluss auf die Prognose analysiert werden. Im Folgenden wird die Prognosegüte des von uns verwendeten Faktormodells in Echtzeit evaluiert sowie die Prognoserevisionszerlegung anhand sog. „News“, welche die Interpretation der Modellprognose vereinfachen kann, erläutert.

Abbildung 1:

Prognoseevaluierung 2016Q1–2015Q4



RMSFE: Wurzel des mittleren quadrierten Prognosefehlers der Modelle im Verhältnis zur Wurzel des mittleren quadrierten Prognosefehlers des Stützzeitraummittelwerts

Quelle: Eigene Berechnungen.

Der der Evaluierung zugrundeliegende Datensatz umfasst etwa 80 monatliche Konjunktur- und Arbeitsmarktindikatoren der Echtzeitdatenbank der Deutschen Bundesbank, knapp 60 umfragebasierte Indikatoren wie den ifo Geschäftsklimaindex oder die Economic Sentiment Indicators der Europäischen Kommission sowie 20 Finanzmarktvariablen (Zinssätze, Aktienmarktindizes, Wechselkurse und Rohstoffpreise). Geschätzt wird das Modell mit bayesianischen Methoden. Es werden Modellvarianten mit zwei, vier, sechs und acht Faktoren geschätzt; darüber hinaus wird eine aus diesen vier Varianten kombinierte Prognose

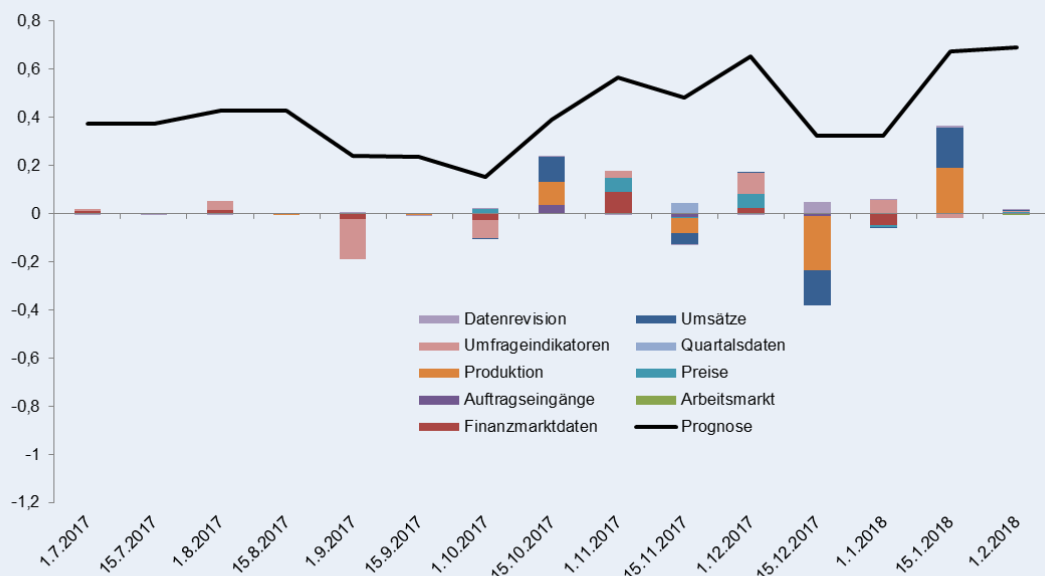
gebildet, wobei alle Modelle das gleiche Gewicht erhalten (sog. „Pooling“).^c Die Prognosegüte des „gepoolten“ Faktormodells wird mit der eines kleinen Modells, das neben der Zuwachsrate des Bruttoinlandsprodukts die Veränderung des ifo Geschäftsklimas und die Veränderungsrate der Industrieproduktion aufnimmt (MF-BVAR), sowie mit den gemäß Reuters-Befragungen erwarteten Zuwachsraten des Bruttoinlandsprodukts verglichen. Als Maß der Prognosegüte wird die Wurzel des mittleren quadrierten Prognosefehlers (RMSFE) der Modelle/Befragungen im Verhältnis zum RMSFE des naiven Modells, welches zur Prognose das Mittel der Zuwachsraten des Bruttoinlandsprodukts im Stützzeitraum verwendet. Geschätzt wird das Modell ab dem Januar des Jahres 1996; der Evaluierungszeitraum reicht vom ersten Quartal des Jahres 2006 bis zum vierten Quartal des Jahres 2015.

Die Prognosegüte der Faktormodelle nimmt ab einem Prognosehorizont von 90 Tagen (im Verhältnis zum Ende des Quartals) kontinuierlich zu und liefert in diesem Zeitraum deutlich treffsicherere Prognose als das Vergleichsmodell. Die Leistung der Faktormodelle ist darüber hinaus vergleichbar mit der der Reuters-Befragungen (Abbildung 1). Eine besonders große Reduktion des Prognosefehlers zeigt sich in der zweiten Hälfte des letzten Monats eines gegebenen Quartals (10 Tage vor Ende des Quartals). Dies dürfte auf die dann erstmals vorliegenden harten Indikatoren wie Industrieproduktion, Auftragseingänge usw. für den ersten Monat des laufenden Quartals zurückzuführen sein.

Die Bedeutung einzelner Variablen für die Prognose lässt sich anhand sogenannter „News“ abbilden. Im Rahmen des Kalman-Filters ist damit die Abweichung des Werts eines neu verfügbaren Indikators von der modellinternen Prognose für diese Variable gemeint. Dadurch wird die Schätzung der Faktoren aktualisiert und dementsprechend auch die Prognose für das Bruttoinlandsprodukt angepasst. Die Revision der Prognose zwischen zwei Datenständen lässt sich dann als gewichtetes Mittel dieser „News“ darstellen (siehe Technischer Anhang). Bei der Bestimmung der Gewichte spielt eine entscheidende Rolle, wie relevant eine Variable ist, sprich wie stark sie mit dem Bruttoinlandsprodukt korreliert. So kann eine nur geringe Abweichung des veröffentlichten Wertes von der modellinternen Prognose für diese Variable zu einer starken Anpassung der Prognose des Bruttoinlandsprodukts führen, wenn die Variable vom Modell als sehr relevant eingestuft wird.

Abbildung 2 :

Zerlegung der Prognoserevision für das vierte Quartal 2017



Linie: Prognose der Zuwachsrate des Bruttoinlandsprodukts. Balken: gewichtete „News“ nach Variablengruppen; die Summe der Balken summiert sich zur Prognoserevision.

Quelle: eigene Berechnungen

Exemplarisch stellen wir dies für die Prognose der Zuwachsrate des Bruttoinlandsprodukts für das vierte Quartals des Jahres 2017 dar (Abbildung 2). Hierzu wurde das Modell mit zwei Faktoren zunächst mit Daten bis zum 15. Juni geschätzt. Danach wurde die Prognose des Modells alle zwei Wochen mit dem jeweils aktuellen Echtzeitdatenstand aktualisiert (Linie). Ebenfalls abgebildet sind die zur Prognoserevision gehörenden gewichteten „News“ (Balken), die zu Variablengruppen zusammengefasst wurden. Insgesamt wies die Modellprognose seit Beginn des zu prognostizierenden Quartals eine steigende Tendenz auf – am 1. November erwartete das Modell einen Anstieg des Bruttoinlandsprodukts von 0,6 Prozent. Aufgrund der schwachen Produktions- und Umsatzentwicklung für Oktober wurde die Prognose jedoch Mitte De-

zember stark nach unten revidiert.^d Nach Veröffentlichung der Novemberwerte im Januar pendelte sich die Prognose jedoch wieder auf ihrem altem Wert ein und lag zwei Wochen vor Veröffentlichung offizieller Zahlen sogar etwas über der amtlichen Meldung der Zuwachsrates des Bruttoinlandsprodukts im vierten Quartal (0,7 bzw. 0,6 Prozent).

Für das laufende Quartal (2018Q1) rechnet das Modell – Stand 10.3. – mit einem Zuwachs des Bruttoinlandsprodukts von 0,5 Prozent. Die Februarwerte der Produktion, Auftragseingänge und Umsätze, die Anfang März veröffentlicht wurden, führten hierbei zu einer spürbaren Revision nach unten (vor der Veröffentlichung dieser Werte lag die Prognose bei 0,9 Prozent). Dabei waren es vorrangig die Umsatzindikatoren, die zu der Abwärtsrevision führten. Auch die Prognose für das zweite Quartal des laufenden Jahres wurde durch die Veröffentlichung jüngster Zahlen nach unten revidiert. Am 15. Februar stand die Prognose noch bei 0,5 Prozent; einen Monat später, also nach Erscheinen der Umfrageindikatoren Ende Februar und der Produktions-, Umsatz- und Auftragseingangsdaten Anfang März, betrug sie noch 0,4 Prozent. Allerdings führten in diesem Fall die Umfrageindikatoren zu der Prognoserevision, während die „News“ der übrigen Variablen vernachlässigbar waren.

^aSchumacher und Breitung (2008) und Jansen et al. (2016) evaluieren Faktormodellprognosen des deutschen Bruttoinlandsprodukts. — ^bDetails des Modells finden sich im Technischen Anhang. — ^cDa die Prognosegüte einzelner Modelle im Zeitverlauf stark schwanken kann und es schwierig ist, in Echtzeit das Modell mit der höchsten Prognosegüte zu identifizieren, führt das Kombinieren mehrerer Modellspezifikationen häufig zu treffsichereren Prognosen (Kuzin et al. 2013; Deutsche Bundesbank 2013). — ^dDer Rückgang der Produktion im Oktober ging auch auf die Lage der Feiertage (Brückentage) zurück, so dass der Rückprall im November erwartet worden war. Solche Einflussfaktoren können von Prognosemodellen typischerweise nicht abgebildet werden.

Literatur

- Deutsche Bundesbank (2013). Prognosemodelle in der kurzfristigen Konjunkturanalyse – Ein Werkstattbericht. *Monatsbericht* (September): 69–84.
- Jansen, W. J., J. Xiaowen und J. M. de Winter (2016). Forecasting and nowcasting real GDP: Comparing statistical models and subjective forecasts. *International Journal of Forecasting* 32 (2): 411–436.
- Kuzin, V., M. Marcellino und C. Schumacher (2013). Pooling versus model selection for nowcasting GDP with many predictors: empirical evidence for six industrialized countries. *Journal of Applied Econometrics* 28 (3): 392–411.
- Schumacher, C., und J. Breitung (2008). Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data. *International Journal of Forecasting* 24 (3): 386–398.

Technischer Anhang

1.1 Faktormodell

Einem Faktormodell liegt die Idee zugrunde, dass der Gleichlauf einer Vielzahl von Variablen durch eine geringe Zahl an nicht beobachtbaren Faktoren zusammengefasst werden kann. Dies geschieht, indem jede Variable in eine gemeinsame (von den Faktoren abhängige) und eine idiosynkratische Komponente zerlegt wird. Für die Dynamik der Faktoren wird dabei häufig eine einfache Vektorautoregression unterstellt. Sei x_t ein $N_m \times 1$ Vektor von stationären Monatsvariablen und f_t ein $R \times 1$ Vektor von Faktoren. Das Modell lässt sich dann anhand der folgenden beiden Gleichungen beschreiben:¹

$$\begin{aligned} x_t &= \lambda_m f_t + u_t; \quad u_t \sim N(\mathbf{0}, \text{diag}(\sigma_1, \dots, \sigma_i, \dots, \sigma_N)) \\ f_t &= \phi f_{t-1} + e_t; \quad e_t \sim N(\mathbf{0}, \Xi) \end{aligned}$$

1.2 Frequenzmischung

Um in dem oben beschriebenen Modellrahmen Quartalsdaten zu berücksichtigen, wird für diese zunächst auf Monatsdatenebene dieselbe Faktorstruktur angenommen. Für die „Monatsrate“ des Bruttoinlandsprodukts gilt also beispielsweise

$$y_{BIP,t}^* = \lambda_{BIP} f_t + u_{BIP,t}; \quad u_{BIP,t} \sim N(\mathbf{0}, \sigma_{BIP}^2)$$

wobei $y_{BIP,t}^*$ nicht beobachtbar ist. Direkt beobachtbar ist lediglich die quartalsdatenbasierte Zuwachsrate des Bruttoinlandsprodukts, $y_{BIP,t}^Q$. Es lässt sich jedoch aufgrund der Zeitreiheneigenschaften des Bruttoinlandsprodukts eine Aggregationsregel finden, die die Quartalswerte mit den unbeobachteten Monatswerten in Verbindung setzt (vgl. Mariano und Murasawa 2005). So gilt approximativ

$$y_{BIP,t}^Q \approx \frac{1}{3} y_{BIP,t}^* + \frac{2}{3} y_{BIP,t-1}^* + \frac{3}{3} y_{BIP,t-2}^* + \frac{2}{3} y_{BIP,t-3}^* + \frac{1}{3} y_{BIP,t-4}^*$$

1.3 Bayesianische Schätzung

Für die bayesianische Schätzung des oben beschriebenen Zustandsraummodells kann auf den Gibbs-Sampler zurückgegriffen werden. Um aus der Posteriorverteilung der Parameter $\theta = \{\lambda_m, \lambda_{BIP}, \Xi, \sigma_{1:N}^2, \sigma_{BIP}^2\}$ und der Zustände $\xi_t = [f_t' u_{BIP,t}]'$ gegeben die Daten – $p(\theta, \xi_{1:T} | x_{1:T_m}, y_{BIP,1:T_q}^Q)$ – zu ziehen, wird auf die leichter zu handhabenden bedingten Verteilungen $p(\theta, | x_{1:T_m}, y_{BIP,1:T_q}^Q, \xi_{1:T_m})$ bzw. $p(\xi_{1:T_m} | x_{1:T_m}, y_{BIP,1:T_q}^Q, \theta)$ zurückgegriffen. Eine Iteration des Gibbs-Samplers zieht dann aus der Posteriorverteilung der Parameter gegeben die Faktoren. Gegeben die Parameterwerte können die unbeobachteten Zustände des Modells mithilfe eines „simulation smoothers“ aus ihrer bedingten Posteriorverteilung gezogen werden (Carter und Kohn 1994 oder Durbin und Koopman 2002). Nach einer Burn-in-Phase konvergieren die Züge aus den bedingten Posteriorverteilungen zu jenen aus der gemeinsamen Verteilung. Die Wahl der Hyperparameter der Priorverteilungen erfolgt im Einklang mit der Literatur zu bayesianischen Faktormodellen (Marcellino et al. 2016).

¹ Aufgrund der Annahme einer diagonalen Kovarianzmatrix von u_t wird die gesamte Kovarianz der Daten durch die Faktoren erklärt (sog. „exaktes“ Faktormodell).

1.4 Prognoserevisionszerlegung anhand von „News“

Innerhalb des Modellrahmens können Prognoserevisionen, die sich aufgrund eines neuen Datenstands ergeben, auf die dafür ursächlichen Variablen zurückgeführt werden. Sei $E[y_{k,t_k}|I_{v+1}] = E[y_{k,t_k}|\Omega_{v+1}] - E[y_{k,t_k}|\Omega_v]$ die Prognoserevision der Variable k zum Zeitpunkt t_k , die sich aufgrund des neuen Datenstands Ω_{v+1} ergibt. $I_{v+1,j} = y_{j,t_j} - E[y_{j,t_j}|\Omega_v]$ sind die „News“ der j -ten neu verfügbaren Variable, sprich die Abweichung der Realisierung von der modellinternen Prognose, gegeben den alten Datenstand. Die Prognoserevision lässt sich dann als gewichtetes Mittel dieser „News“ darstellen: $E[y_{k,t_k}|I_{v+1}] = B_{v+1} I_{v+1} = E[y_{k,t_k} I'_{v+1}] E[I_{v+1} I'_{v+1}]^{-1} I_{v+1}$, wobei die Terme $E[y_{k,t_k} I'_{v+1}]$ und $E[I_{v+1} I'_{v+1}]^{-1}$ mithilfe des Kalman-Glätters bestimmt werden können (vgl. Banbura et al. 2013).